Visual Analytics in Biomedical Applications

Paul Parsons[1], Robert Mercer[1], Kamran Sedig[1,2], Peter Rogan[1,3], Joan Knoll[3], Maryam Khordad[1], Jon Demelo[1]

1 Department of Computer Science; 2 Faculty of Information and Media Studies; 3 Schulich School of Medicine & Dentistry

NATURAL LANGUAGE PROCESSING

In Natural Language Processing (NLP) research, computational techniques are used to extract useful information from natural language text

Typical NLP tasks are named entity recognition, parsing, part-of-speech tagging, sentence breaking, concept normalization, and relationship extraction

NLP of biomedical text is a very active research area, and presents its own unique challenges for extracting useful information about biomedical entities and relationships

We are engaged in research toward the end of identifying biomedical entities in scientific articles (e.g., genes, phenotypes, proteins), as well as identifying implicit relations that exist within the natural language text (e.g., genotype-phenotype relations)

INDEXING AND SEARCH

Biomedical researchers and clinicians often need to consult and work with scientific articles

There are approximately 24 million articles in PubMed. For any given biomedical activity, there may be thousands of relevant articles

In addition to NLP, we are indexing all PubMed citations (metadata) to enable fast and accurate search of biomedical entities and relations

During indexing, we perform natural language processing to identify important concepts, such as gene and phenotype names. This is not a trivial task, as many concepts (e.g., genes) are written in a variety of ways. We are also finding implicit relationships in the text, such as genotype-phenotype relationships

We are using IBM Watson technology, as well as popular open-source products, for indexing and search. Users can search for terms and retrieve results in a matter of seconds. Results are then visually represented for the users to work with

FURTHER READING

Sedig, K. & Parsons, P. (2013). Interaction design for complex cognitive activities with visual representations: A pattern-based approach. AIS Transactions on Human-Computer Interaction, 5(2), 84–133.

Parsons, P., Sedig, K., Didandeh, A., & Khosravi, A. (2015). Interactivity in Visual Analytics: Use of Conceptual Frameworks to Support Human-Centered Design of a Decision-Support Tool. 48th Annual Hawaii International Conference on System Sciences (HICSS). IEEE.

Khordad, M., Mercer, R. E., & Rogan, P. (2011). Improving phenotype name recognition. In Advances in Artificial Intelligence (pp. 246-257). Springer Berlin Heidelberg.

INTRODUCTION

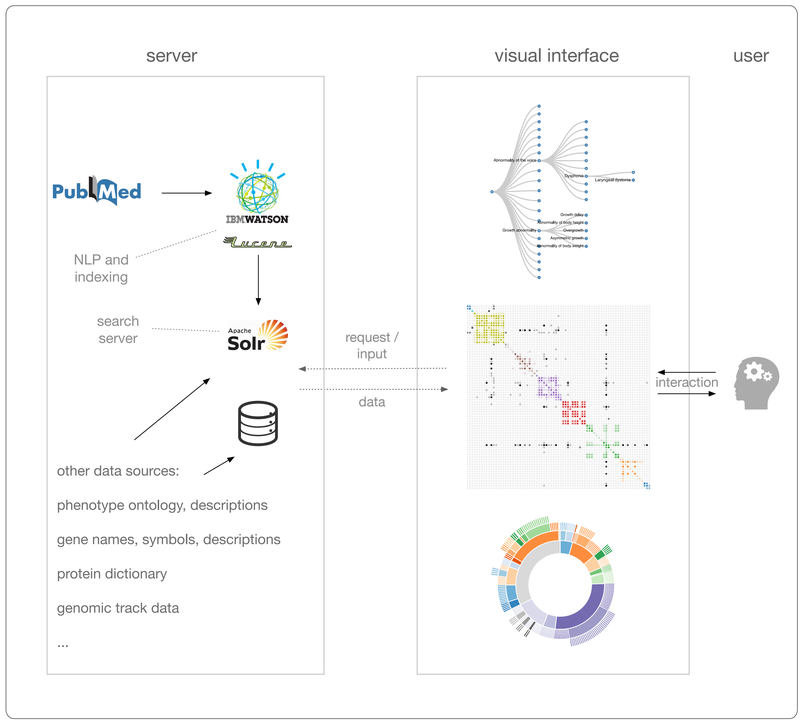
Biomedical tasks often require working with very large and complex information spaces (e.g., scientific articles, gene databases, phenotype ontologies, microarray data, and others)

It can be very challenging and time-consuming to navigate through such a space in order to find and make use of the relevant information for a given activity

Visual analytics combines the strengths of computers and humans to support information-intensive complex activities. Interactive visual representations are employed to amplify perceptual and cognitive tasks of users

We are developing an interactive visual analytics system that supports the complex activities of biomedical researchers and practitioners. On the computational side, we are engaged in natural language processing research, in an attempt to extract useful information from millions of scientific articles. On the human side, we are engaged in research on visual representation design and interactive cognition, to best support users as they work with the information

SYSTEM OVERVIEW

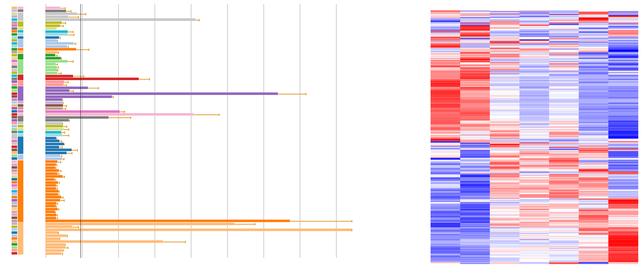


VISUAL REPRESENTATIONS

The human visual system has a higher bandwidth than any other perceptual system

Unlike descriptive representations of information (e.g., natural language, mathematical formulae), depictive visual representations (VRs) can be perceived in a parallel fashion, which enables much information processing to be offloaded onto the visual system, decreasing cognitive load while working with information

VRs, especially if they are interactive, can greatly assist in understanding and working with complex information spaces



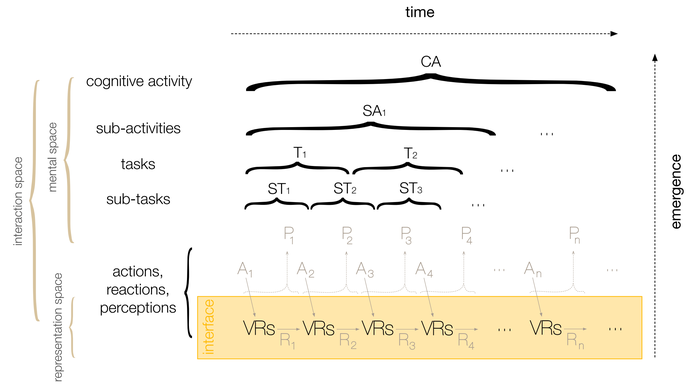
INTERACTIVE COGNITION

Research has shown that our cognition is distributed across our brain and external environment

We carry out cognitive activities (e.g., problem solving, decision making) largely through interacting with external information

Interactive visual representations support complex tasks, such as those performed by biomedical clinicians and researchers, by combining interactive cognition and visual reasoning

When visual analytics systems are designed well, they support the emergence of complex activities over time



Cognitive Engineering Lab